
Editor's message: Groundwater modeling fantasies —part 1, adrift in the details

Clifford I. Voss

Fools ignore complexity. Pragmatists suffer it. Some can avoid it. Geniuses remove it. ...Simplicity does not precede complexity, but follows it. (Epigrams in Programming by Alan Perlis, a computer scientist; Perlis 1982).

A doctoral student creating a groundwater model of a regional aquifer put individual circular regions around data points where he had hydraulic head measurements, so that each region's parameter values could be adjusted to get perfect fit with the measurement at that point. Nearly every measurement point had its own parameter-value region. After calibration, the student was satisfied because his model correctly reproduced all of his data. Did he really get the true field values of parameters in this manner? Did this approach result in a realistic, meaningful and useful groundwater model?—truly doubtful. Is this story a sign of a common style of educating hydrogeology students these days? Where this is the case, major changes are needed to add back 'common-sense hydrogeology' to the curriculum. Worse, this type of modeling approach has become an industry trend in application of groundwater models to real systems, encouraged by the advent of automatic model calibration software that has no problem providing numbers for as many parameter value estimates as desired. Just because a computer program can easily create such values does not mean that they are in any sense useful—but unquestioning practitioners are happy to follow such software developments, perhaps because of an implied promise that highly parameterized models, here referred to as 'complex', are somehow superior. This and

other fallacies are implicit in groundwater modeling studies, most usually not acknowledged when presenting results. This two-part Editor's Message deals with the state of groundwater modeling: part 1 (here) focuses on problems and part 2 (Voss 2011) on prospects.

Beliefs

Some modelers believe that the more spatially complex the model, the more details of field data it reproduces, the better a representation of reality it becomes and the better a tool it becomes. Sometimes this approach is justified by saying that details are needed because real geology has a complex appearance. The most detailed picture of a region available from modern surface, borehole and geophysical data might include many mapped fracture zones, several units containing different geologic fabrics possibly folded and/or offset vertically, perhaps intruded lithology that crosses bedding, perhaps surfaces separating some units unconformably, a variety of soils, and a complex pattern of surface-water bodies and groundwater recharge. It is a typical occurrence that these 'detail-modelers' and also some geologists and other practitioners whose focus is on collecting and interpreting field data, strongly insist that all minutiae of geologic structures they have discovered must be included in the model, or the model will be neither complete nor correct. This indicates a lack of perspective on the functioning of hydrogeologic systems, as they do not realize that only some of their details are important to modeled system behavior and to the questions to be answered by the analysis. They need to understand that effective groundwater modeling involves discovery of which these factors might be. It is also clear that no hydrologic data set in the world can provide sufficient basis for estimating parameters of all of the features in such complex hydrogeologic models. Surely, the detailed picture is valuable as it provides one particular type of earth understanding, but one that should not be imported wholesale into a groundwater model; rather it should be used to help inform the modeling process.

A spatially complex model such as just described, is necessarily also a highly parameterized model. Another type of highly parameterized model is one that may have relatively simple geologic structure, but highly detailed

Received: 16 September 2011 / Accepted: 23 September 2011
Published online: 25 October 2011

© Springer-Verlag (outside the USA) 2011

Clifford Voss is the Executive Editor of *Hydrogeology Journal*.

C. I. Voss (✉)
US Geological Survey,
345 Middlefield Rd, MS 496, Menlo Park, CA 94025, USA
e-mail: cvoss@usgs.gov

spatial variations in model parameter values (usually in hydraulic conductivity) to allow the model to achieve good fit to available data.

Is there any substance to the belief that spatially complex or other highly parameterized models are better? No proof is ever given of the added value of highly parameterized models. In fact, such groundwater models are no better representations of subsurface systems than arbitrary fitted functions with many parameters. For example, if we have N observations of a field parameter such as hydraulic head in a subsurface system, we can create a polynomial of order $(N-1)$, or a spline or many other arbitrary functions with N parameters, whose values can be adjusted to allow the function to fit all of these data perfectly and to predict values between data points and outside of the data range. Such functions make no claim to be mathematical representations of any of the physical processes of groundwater flow. After parameter-value estimations for this perfectly fitted black-box model, we can make projections of hydraulic head for the subsurface system for which it was created—and as with all fitted models, the correctness of the black-box projection decreases with time and distance from the measurement points. Both the black-box model and a highly parameterized groundwater model with a similar number of fitted parameters provide equally good (or poor) projections, because the highly parameterized model's modeled physics makes little difference to the model's ability to fit the data. The equation describing groundwater-flow physics solved by the highly parameterized groundwater model is being used only for interpolation, and there is no guarantee that the many values of estimated parameters correctly represent the subsurface properties. In contrast, a groundwater model with very few parameters whose values are adjusted to allow the model to approximately fit the important parts of the same available data manages this fit because it relies on the calculated physics of groundwater flow. Was not this the reason we used a groundwater model for the analysis, rather than an arbitrary function, in the first place?

Clearly geology can be studied and understood as well as possible for its own sake, including whatever geologic details are needed. However, just because real geology 'looks' complicated does not mean that we must bring all of that complexity into a descriptive model of the groundwater system. In the cases where this is done, the unavoidable result is a complex fitted model that contains its many details arranged in false spatial patterns and with wrong parameter values. Simply put—this groundwater model is not useful.

A strong belief in the correctness or superiority of a complex model because all available details have been included is unfounded and unprovable. This requires faith—and faith is a human behavior that does not belong as a main component in informing the daily practice of managing water resources. Highly detailed models are fantasies.

Problems with complex models

Groundwater modeling is by definition an implicitly unsystematic exercise that, in its most useful form, is

based on judgment and scientific intuition. Parameter values in a model are often estimated so that modeled behavior fits measurements, with the intent of making the model a good representation of the groundwater system. Estimation had, in the past, been done manually by adjusting model parameter values until a good fit has been obtained. Recently available automatic-estimation software surely makes this part of modeling work easier.

The primary technical problem underlying all parameter estimation is that estimating many parameters at once inevitably results in correlation. Anyone attempting automatic estimation, using one of the software tools available, immediately finds from the output of the estimation program, that many of the target parameter estimates are correlated. Correlation means that the estimated values of correlated parameters can each be adjusted up or down to achieve the same degree of fit of simulated values to observed values, so that it is impossible to separately determine individual parameter estimates from the data available. Only the value of one of each pair of correlated parameters or of the appropriate function of these can be determined. The estimation of correlated parameters, as is routinely done in groundwater modeling today, likely provides the wrong values of these parameters. Nevertheless, model projections are then routinely made using a model populated with the wrong parameter values.

The primary philosophical problem in groundwater modeling is that when creating a model, some assumptions made about the structure of the subsurface and about external hydrologic processes that impact the subsurface are inevitably going to be wrong (and what exactly is wrong is not ever likely to be discovered). Thus, the created model does not represent the important subsurface features or hydrogeologic factors. Such incorrect assumptions lead to estimation of wrong parameter values even if the assumed spatial pattern of parameters (zonation) happens to be correct—and even if, miraculously, the target parameter estimates are not correlated. However, guessing the true zonation correctly is also not likely. (For example, it may be well known that a confining unit exists and can be confidently designated as a single model zone, but dividing it into many separate zones, each with a different parameter value, is not usually feasible.) An error in zonation, assumed model structure, or in some value assumed for input parameters, will cause automatic fitting to generate errors in other parameter values. These erroneous values may be organized in a realistic-appearing spatial trend that some modelers naively accept as reality. How can reality of a trend or newly discovered model parameter zone be determined without further targeted collection of field data? An additional round of field work is a possibility that is not available in most projects. So what is the value of discovering such a pattern, and how should its uncertainty be related to those interested in how the model answers the questions at hand?

A complex model with many parameters and many details may fit the existing data perfectly well—but stress it in a new way, e.g. add a pumping well in a new

location, and the wrong result will be obtained for response at that point. The many-parameter black-box function that is devoid of groundwater physics can be expected to do equally well (in other words, equally poorly) for this problem.

The primary thinking required for detailed data fitting of highly parameterized models is merely how to push buttons and adjust knobs on computer programs so that they run as effectively as possible toward obtaining detailed fits to data. This time-consuming and tedious process is the wrong objective for modeling. Too much time is spent on it in many projects, creating a deficit of useful model-based results. Instead, a better objective is learning something from a model-based analysis. This should be the primary goal of all groundwater modeling. Learning rather than fitting can also be a more personally satisfying task for the hydrologist.

Misleading bases for modeling

Use of other types of measurements besides hydraulic heads to help constrain groundwater-model fitting, or use of statistical approaches, in the face of non-uniqueness and ambiguity of typical results, are not as useful as fixes as may appear at first glance.

Baseflow. Measurements of groundwater discharge, for example, baseflow, do not provide a foolproof escape from non-uniqueness. Actual stream baseflow depends strongly on the hydraulic conductivity distribution immediately around the streambed. This is rarely well known and is usually not represented well in a model. Without the correct local hydraulic conductivity distribution, the groundwater discharge is wrongly simulated, forcing other model parameters to compensate during automatic estimation by adjusting to wrong values. It is sadly also not possible to simultaneously estimate flow and hydraulic conductivity at the same location because their impact on simulated hydraulic heads is strongly correlated by the physics. The result is that flow observations can sometimes help to inform model construction, but not necessarily more than head observations.

Concentration and temperature. Use of field data on solute concentration or temperature, in addition to measured hydraulic heads and flows, has been touted as a means to help constrain estimates of hydrologic model parameters. This assumption is often fallacious because heads are created by one set of physical processes in heterogeneous geologic fabrics, whereas concentrations and temperatures are created by different sets of processes. Heads diffuse through the subsurface. Solutes and heat also diffuse but are additionally transported by flowing groundwater—via advection. Diffusing changes in head propagate through all of the subsurface fabric, through low- and high-permeability patches, more slowly through patches of low hydraulic diffusivity. Advection occurs primarily through connected patches of the more-permeable fabric. Where there is active flow, actual subsurface concentrations/temperatures are therefore determined by

only a particular subset of the geologic fabric. Consequently, head measurements and concentration/temperature measurements have been created by different processes that have sampled different parts of the subsurface—and so these cannot be simultaneously used to define parameters of the same model. In particular, the hydraulic conductivity value that would be determined from fitting to measured heads alone is quite different from the value determined from fitting measured concentrations alone. Indeed, basic understanding of flow and transport processes that occur in heterogeneous geologic fabrics would lead us to expect that these estimated values should be different. Finally, there is little value in fitting to some weighted combination of observed heads and concentrations, as allowed by available software, because there is one parameter estimate available from heads and a different one from concentrations. Weighting only selects an arbitrary intermediate value, depending on the practitioner's arbitrary selection of a weight. It is useful to recognize that obtaining the extreme values does help us to better understand the subsurface situation; however, the only setting in which using both heads and concentrations simultaneously to define groundwater model parameters makes sense is a perfectly homogeneous aquifer where both propagation processes sample the same parts of the geologic fabric—all of it. There are not many aquifers of this sort. In general, heads and concentrations should not be used simultaneously for parameter estimation when constructing a groundwater model. A similar difficulty impacts the simultaneous use of interpreted (e.g. isotopic) water age and heads for calibration.

Recharge. An external factor most-often used as the primary driver of modeled groundwater flow (as a source-type boundary condition) is the amount of groundwater recharge. Regrettably, groundwater recharge is rarely well defined. It is not unusual that the quantity typically measured, the amount of water that passes from the air or from a surface-water body through the ground surface, is patently uninteresting in the water budget of the aquifer below. For example, where a water table is at or near the ground surface, and ground-surface topography has high frequency of lateral elevation variation, much of the water that passes downward through the ground surface discharges close to its entry point, following very short paths between entering and departing the subsurface. Similarly, after a field hydrologist measures the amount of water passing into the ground surface, much of this water may discharge to a nearby surface-water body via a thin surficial permeable layer and does not reach the deeper zone being studied. In these situations, is the total flux into the ground the practical useful value of groundwater recharge? No, only the flux that enters the depth of interest, usually an aquifer, is a useful quantity. Even for the water-table aquifer in the aforementioned example, a high total flux initially enters, but the practical recharge value is the net flux in (subtracting what quickly discharges to the rough ground surface). In addition, recharge may not be spatially uniform due to patchiness in amount and location of vertical move-

ment of groundwater to a water table or through a confining unit to a confined aquifer. Recharge values may not be constant, changing with time and location as a result of groundwater processes. Clearly, recharge is not easily measured in the field, and values so-determined may be quite different from what actually enters an aquifer and from what a model would indicate based on known values of heads and hydraulic parameters. Thus, a useful recharge value might best be obtained indirectly, as a result of a modeling analysis based on other hydrologic data such as head or groundwater age. This reverse approach may be preferable to employing a possibly wrong, externally determined value of recharge as the main driver of flow in the modeled system, as is commonly done.

Statistics. Directly facing the difficulties entailed by a single model representation of the complex subsurface, where that single assumed representation is presumably wrong, statistical methods have been proposed to consider an ensemble of possible models or subsurface patterns and parameter values. Statistical approaches have the apparent advantage of not needing to specify a particular subsurface structure, only the statistics of structures and values. Complex spatial distributions of hydrogeologic properties can be generated using statistical models with only a few parameters. The limitations of such approaches are several. (1) The scale of variability of some features may be larger than the spatial region of the process being predicted, bringing into question the applicability of a simple statistical distribution for the region, in which perhaps a single feature and its location and properties (such as an individual fracture) may control the behavior of importance. The pragmatic objective might then be to rather find and define the feature, than to describe its presence statistically. (2) Defining the statistical distribution in the field perhaps requires even more data for reliable definition than would a deterministic model of subsurface structure in the same area. The great intensity of field data required to develop the actual statistical distribution will never be collected on more than a handful of demonstration sites worldwide. (3) Even where such data are collected, the parameters and type of statistical distribution may not remain constant in space—and quantification of such variation would require even more measurements. (4) The particular location being studied has only a single structure, and so may exhibit ground-

water behavior that significantly deviates from a mean prediction, bringing into question the value of providing a mean and variance for the process and location in question. How can managers use such statistics practically? (5) The statistics may not include parameterization of all necessary aspects of spatial distributions (e.g. connectivity) that control the groundwater behavior of importance to the questions being asked. (6) Even if there were a practical way for managers to use the predicted mean and variance, the statistical result is itself highly uncertain, because it is based on assumed probability distributions that are most often not based on actual field data and that are themselves highly uncertain for all the previously given reasons.

The discussion concludes in part 2 (Voss 2011). This two-part Editor's Message focuses on standard saturated-zone groundwater-flow modeling, not on solute or energy transport modeling or on unsaturated zone modeling, all of which are subject to concerns at least as great. An analogous discussion on groundwater modeling, published 13 years ago, with interesting quotes on quantitative methods and simplicity by M. King Hubbert and M. Biot, may be found in Voss (1998).

If you can't reduce a difficult engineering problem to just one 8½×11-inch sheet of paper, you will probably never understand it. (Ralph Brazelton Peck, a soil mechanics engineer, quoted in DiBiagio and Flaate 2000).

References

- DiBiagio E, Flaate K (2000) Ralph B. Peck: Engineer, Educator, A Man of Judgement. Publication No. 207, Norwegian Geotechnical Institute (NGI), pp 73. <http://peck.geoengineer.org/books.html>. Cited September 2011
- Perlis A (1982) Epigrams in programming. SIGPLAN Notices 17 (9):7–13
- Voss CI (1998) Editor's Message - Groundwater modeling: Simply powerful. Hydrogeol J 6(4):A4–A6
- Voss CI (2011) Editor's message: Groundwater modeling fantasies—part 2, down to earth. Hydrogeol J. doi:10.1007/s10040-011-0790-6